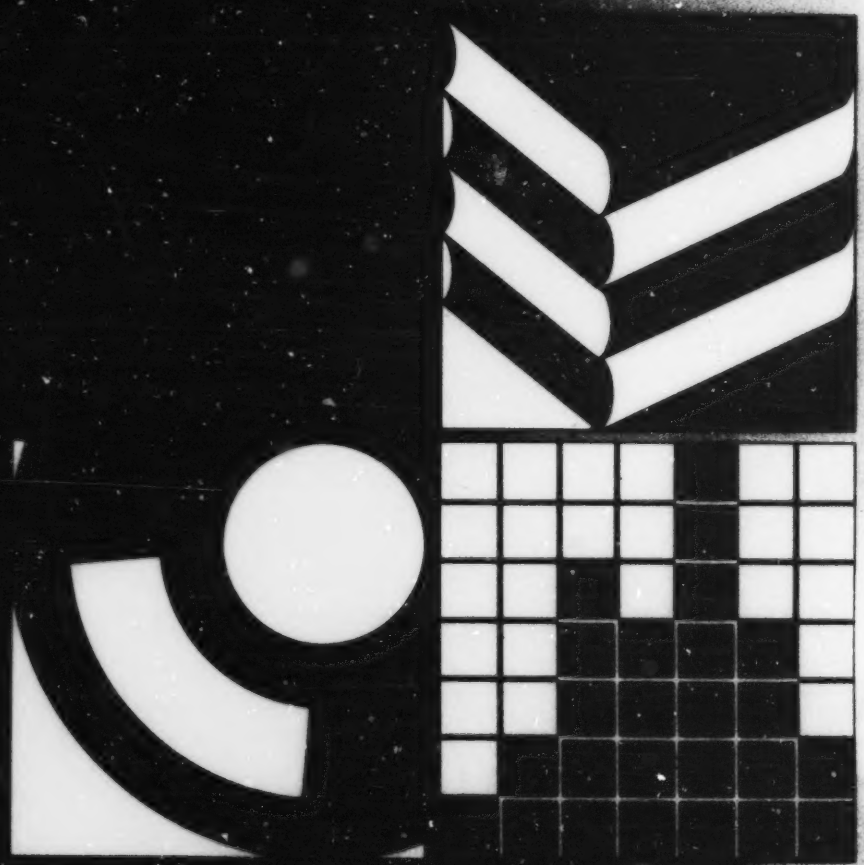
PHOCUS



Canada

-4420/

PHOCUS

A BIBLIOGRAPHIC DATA BASE

FOR THE

CONSERVATION OF PHOTOGRAPHIC MATERIALS

by

Klaus B. Hendriks

and

Diane Hopkins

Public Archives Canada Archives publiques Canada

Canadian Cataloguing in Publication Data
Public Archives Canada.
PHOCUS, a bibliographic data base for the conservation of photographic materials
Text in English and French with French text on inverted pages.
Title on added t.p.: PHOCUS, base de données bibliographiques pour la conservation des documents photographiques.
DSS cat. no. SA82-3/1-1985
ISBN 0-662-53453-0
1. PHOCUS (Information retrieval system). 2. Information storage and retrieval systems--Photographs--Conservation and restoration.
I. Hendriks, Klaus, B. II. Hopkins, Diane. III. Title. IV. Title: PHOCUS, base de données bibliographiques pour la conservation des documents photographiques.

Z699.5.P46P8 1985 025'.0677'028 C85-097000-8E

The PAC intends, in future years, to publish reports of general interest on a variety of archival techniques.

© Minister of Supply and Services Canada 1985
Cat. No. SA82-3/1-1985
ISBN: 0-662-53453-0

TABLE OF CONTENTS

FOREWORD v

1 INTRODUCTION 1

2 BACKGROUND 1

- The Collection..... 1
- System Objectives..... 2
- Selection of MINISIS..... 3

3 DEVELOPMENT OF PHOCUS 4

- Online Thesaurus..... 5
- UDCEXIT: A Processing Routine for the Universal Decimal Classification..... 5
- Recursive Record Structure..... 5
- Programmable Function Keys..... 5
- User Views to De-code and Filter Data..... 6
- Security Profiles also Measure Performance..... 6

4 CONCLUSION 6

- PHOCUS Today..... 6
- Future Developments..... 7

APPENDIX A : Extract from Functional Specifications and System Design Manual 9

APPENDIX B : Figures and Table..... 13

FOREWORD

The processed photographic layer remains a nearly ideal medium for storing information in a small space. Practical applications of this medium include microfilms and fiches, historical still photographic negatives and prints, motion-picture films, aerographic films, industrial and medical X-ray films, spectroscopic plates, duplicating and copying films and others. They may be in either black-and-white or colour. The permanence of these records, and procedures for their preservation, concern both manufacturers and users. For the past nine years, the Picture Conservation Division of the Public Archives of Canada has collected a considerable amount of published information concerning the manufacture of light-sensitive materials, the properties of processed photographic films and papers in terms of their stability, their preservation through control of environmental conditions and their restoration in those instances where damage has already occurred. This literature has been useful in clarifying the factors that determine the longevity of processed photographic materials and in planning and executing laboratory experiments.

An automated data base has been developed in order to thoroughly index these sources of information through a multitude of access points. It is known as PHOCUS. While data entry is continuing, PHOCUS has been shown to function well in demonstrations at the Public Archives of Canada, as well as in Los Angeles (California), via telephone hook-up to Ottawa, during the annual conference of the American Institute for Conservation (AIC). Although the main thrust of the data base's contents are processed photographic materials, PHOCUS contains several hundred references to related materials, particularly paper - that classical record material of libraries and archives - and numerous recent bibliographic citations on the preservation of magnetically recorded information, emerging video disc and digital optical disc technologies.

The purpose of this report is to summarize the planning and development of the PHOCUS data base. As it is unique in the field of conservation of cultural heritage, it may well serve as an example for future data bases in related fields.

Klaus B. Hendriks
Director
Picture Conservation Division

1 INTRODUCTION

The Picture Conservation Division of the Public Archives of Canada (PAC) has a unique collection of technical documentation on the conservation of photographic materials. Access to this collection has recently been automated by the creation of an in-house information storage and retrieval system, using MINISIS. Both the system, and the bibliographic data base it supports, are referred to as PHOCUS:

PHO tograph **C** onservation **U** niversal Decimal Classification **S** ystem.

The establishment of this system will provide the Division with several benefits:

- circulation control of the items in the collection;

- automated updating of the classification scheme and thesaurus, and subsequent indexing amendments that such changes will entail;

- additional access to the documentation via date, source of publication, language, country of origin and any other data elements which may be required at a future time;

- breaking up a Universal Decimal Classification expression into its component parts, so that it may be inverted for fast access searching, with or without thesaurus support.

Because of the uniqueness of this collection, the Public Archives of Canada plans to make the PHOCUS data base available to specialized research centres through online access. Alternatively, in response to requests from such centres, copies of the data base may be made available on magnetic tape. The Picture Conservation Division will also publish technical bibliographies from the PHOCUS data base for general distribution.

2 BACKGROUND

The Collection

Over the years, in support of the Division's activities in the field of photograph conservation, a collection of approximately 5,000 published and unpublished documents has been accumulated. It is a comprehensive collection of international scope, covering the following subject areas:

- history of photographic processes and materials;
- selected aspects of photographic science and technology;
- manufacture of photographic plates, films and papers;
- photochemistry as related to photography;
- permanence of processed photographic materials; and
- conservation of component materials.

Presently, the largest concentration of items covers black-and-white silver gelatin processes.

The collection is classified according to the Universal Decimal Classification (UDC), and indexed by a thesaurus. This thesaurus, which has been developed in-house, is linked to the classification scheme. In the course of using the UDC, the Division submitted a number of proposals to update the schedule for photography. These changes were finally approved by the Fédération Internationale de Documentation (FID), governing body of the scheme, in January 1984[1].

Both the classification scheme and thesaurus were created to provide detailed subject access to this collection through a conventional card catalogue.

System Objectives

This manual system proved to be an inefficient use of the librarian's time as there were no clerical staff available to file the catalogue cards (an average of 6 cards per document indexed). The users found that subject and author access to this collection was not sufficient to retrieve items that were remembered only by date or journal of publication. The production of subject bibliographies was also slow. When permission to automate was granted, the following system objectives were identified in order to improve upon the weaknesses of the manual system[2].

The system was planned to:

1) accommodate a maximum of 10,000 records;
2) provide an increased number of searchable fields (15 minimum);
3) improve response time for searching and processing;
4) store abstracts and allow free-text searching;
5) reduce the cost of processing new records;
6) reduce the unit cost for retrieval;
7) reduce the cost of producing bibliographies;
8) allow flexibility in searching and system design;
9) permit online authority control;
10) make future non-bibliographic data applications possible;
11) provide additional hard copy products; and
12) permit remote communication.

---

[1] Notice of the amendments was published in Extensions and Corrections to the UDC 1983. Series 12, No. 1: 19-23; May 1984, (FID Publication 248/12:1).

[2] Hopkins, Diane. Is an Automated In-House Information Retrieval System a Cost-Effective Means of Providing Subject Access to a Special Collection? London: School of Library and Information Science, University of Western Ontario, 1981. pp. 10-11.

Selection of MINISIS

The initial feasibility study[3] indicated that a microcomputer system with hard disc storage would be the most cost-effective means of achieving the system objectives. Since the study was conducted, the appearance of data base management systems (DBMS) for microcomputers makes this option even more attractive for future applications. At the time, however, no particular product seemed appropriate. The Division investigated the possibility of participating in the joint development of a new package with the Institute of Machine Documentation, under the sponsorship of UNESCO. Although that proposal was not followed through, results of the project should be available some time this year,[4] and are also worthy of consideration by those contemplating creating new bibliographic data bases.

Almost by default, the Division was eventually given the opportunity to develop its application on a configuration not previously considered during the initial feasibility study. Another division of the PAC had just concluded negotiations for a facilities management contract. This meant that a flat fee would be paid for use of an entire minicomputer at a service bureau, regardless of connect time, CPU time, or disc space actually being used. Excess capacity on the system permitted us to develop our application on the same computer at a cost far less than standard time-sharing rates. There was a choice of two DBMS packages already available on the system: IMAGE, the Hewlett-Packard product resident on the system, or MINISIS, an integrated software package developed especially for bibliographic applications. As PAC was already a sub-licensee of MINISIS,[5] there was no need to acquire a copy of it. MINISIS was chosen over IMAGE for the same reasons that the original MINISIS development team rejected IMAGE:[6]

- capability to handle true variable-length records;
- capability to handle variable length, variably occurring fields;
- capability to handle subfields;
- capability to extract search keys embedded in text strings; and
- capability to handle long, descriptive abstracts.

---

[3] Ibid., p.35.

[4] Anon. "Information Technology Projects", General Information Programme Unisist Newsletter. Vol. 11 no. 2, 1983. pp. 24-25.

[5] Presently, MINISIS is marketed in North America exclusively by Systemhouse Ltd.

[6] Daneliuk, Faye A. Information Retrieval and Library Management: An Interactive Minicomputer System. Ottawa: International Development Research Centre, 1978, 16 p. (IDRC-TS14e).

The above features are all requisites of sophisticated bibliographic applications. Other features that make MINISIS generally applicable to any text-based system are:

- modularity: all sub-routines are independent, allowing easier maintenance and extension;

- independence of data: the software is stored separately from actual data input by an application, making it possible for different applications to share the same software without sharing data bases;

- flexibility: capable of handling data in almost any physical form; provides a wide variety of outputs, permits redefinition of data after data entry has already begun;

- user-friendliness: searchers and infrequent users will find the QUERY language relatively easy to learn and use;

- mission orientation: capability to accept output from other information systems; viable within a small organization; and

- cost-effectiveness: new applications can be designed, tested and implemented in 6-12 months.

More specific features of MINISIS, which were found to be attractive at the selection stage, are described in the next section of this paper.

3 DEVELOPMENT OF PHOCUS

The person who assumed the role of Data Base Manager participated fully in the design and testing phases, and is responsible for the ongoing maintenance of the application. This learning experience tended to extend the development period for the application, but ensured a thorough knowledge of the complexities of the design. It should be noted that, contrary to the sales literature, MINISIS is not easy to use at any level beyond data entry without thorough training. Its inherent flexibility increases the number of design decisions, and hence the potential complexity of a particular application design. However, a person with only elementary programming skills, and an intimate knowledge of the application, should be able to fill the Data Base Manager's role successfully after completing the necessary MINISIS training.

Detailed functional specifications were derived from the system objectives previously mentioned (see Appendix A). These specifications were then translated into a data base and application program design which employs MINISIS features, and respects its limitations. This is the stage at which the knowledge of an experienced analyst/programmer becomes invaluable[7]. While the entire system cannot be fully described in this report, the following features of the Division's application may be of particular interest to other data base creators:

---

[7] Bill Page served as analyst/programmer on this project, which was contracted to Infoman, Inc. (Contract No. 0E283-07009)

Online Thesaurus

The faceted UDC expression functions as a type of shorthand (see Figure 1). Each facet in the expression has an English language equivalent in the online thesaurus, so that retrieval is possible using terms from either language. Translation is automatic. The multilingual MINISIS thesaurus structure also supports relational searching, i.e., broader (BT), narrower (NT), related (RT), any (ANY), and forbidden terms (USE). Examples of such relationships are shown in Figure 2.

UDCEXIT[8]: A Processing Routine for the Universal Decimal Classification

A special processing routine, called UDCEXIT, was written to analyse the UDC expression, break it into its component facets, and store these as descriptors in the master record. This saves time during data entry, since equivalent English terms do not have to be typed. Also, some saving in disc space is derived as UDC facets are generally shorter than English descriptors. Figure 3, shows the transformation of the UDC expression from a single character string to a set of English language descriptors. The actual descriptors (UDCKEYs), which are generated and stored by UDCEXIT, are automatically translated into their English language equivalents by the online thesaurus for display in the complete bibliographic record.

Recursive Record Structure

Figure 4 shows a schematic view of the recursive record structure developed for PHOCUS. This structure minimizes duplication of data during data entry, and reduces redundancy of data stored. This is particularly appropriate for a collection where numerous versions of the same document are collected, (one content: several instances) and where numerous items all reside in the same source document (several child records: one parent record). This structure conveniently accommodates the description of 10 different record types (see Figure 5). For search purposes, a user view of these files is defined, which links all component parts into one master bibliographic record. The underlying structure is thus transparent to users at this level (see Figure 3).

Programmable Function Keys

User programmable function keys on a terminal are a valuable time-saving feature. Our Televideo 950 terminal has 11 keys which can be programmed to deliver a character string to the computer that will be interpreted as if it had been typed in whole by the operator. In order for the keys to contain these lists, they must be stored in a program on the host computer, and down-loaded into the keys when the operator logs on to MINISIS. Otherwise, the operator would have to manually program each key every time the power is switched on.

---

8 The UDCEXIT module will be submitted to the international User-contributed Library for distribution with future versions of MINISIS.

Each record type shown in Figure 5 requires a different subset of data elements in order to complete the bibliographic record. Each list of necessary fields is stored in one of these function keys (see Figure 6). The operator simply hits the key which designates the record type, and the appropriate fields are then prompted automatically.

User Views to De-code and Filter Data

The essence of data base management systems is that they permit data, which is captured only once, to be repackaged into seemingly different data bases. By defining a particular user view, the Data Base Manager can filter only the fields required from the master record, and combine encoded fields with look-up tables to provide output with full names rather than the code which was stored in the master record. In addition, MINISIS provides an initial restriction feature that can be used to segregate subsets of records in a data base. For example, the master file can be divided into confidential and public record groups.

Security Profiles also Measure Performance

Five basic security profiles have been identified for this system: Data Base Manager, Librarian, Data Entry Operator, Staff, Others. Each profile limits the user to certain data bases only, as well as the functions which he/she may perform on any data base. Sets of user names within these levels permit collection of performance measurement statistics, because each session is logged by the computer's resident accounting system. For example, at the Librarian level, DEVTHES and MODIFY distinguish work related to thesaurus development from checking and updating of records created by the data entry operator.

4 CONCLUSION

PHOCUS Today

Retrospective conversion of approximately 2,500 manual records has been completed. The Data Base Manager found that substantial maintenance work, especially related to the thesaurus and print formats, has been necessary. A backlog of 2,500 documents is presently being catalogued, and the collection continues to grow at the rate of 1,000 items per year. Table 1 summarizes projected cost and time estimates, and actual costs and time spent.

At present, little effort is devoted to diffusion. However, the Division is considering a request by the Photo Materials Group of the American Institute for Conservation of Historic and Artistic Works to publish current references in the AIC Newletter. This would be done as a by-product of our in-house current awareness bulletins. Enquiries are dealt with as expediently as possible, but the staffing situation precludes the encouragement of outside users at this time.

Future Developments

The classification scheme is of particular interest to those whose field is photograph conservation. The British Standards Institution is presently preparing a revised edition of the schedule for photography, which will incorporate all the amendments introduced into the scheme, and will serve as an international standard for the organization of this literature[9].

Concerning the automated system itself, the following enhancements are planned:

- Modification of the thesaurus module design to permit output of alternate language equivalents as descriptors.

- Completion of the module to provide circulation control of the collection. This work was started during the development phase, but requires elaboration and testing before it can be fully implemented.

- Design of a print format to print Interlibrary Loan request forms, as required by the Public Archives Library (PAL). This will eliminate duplication of effort between the original request to acquire an item and its actual cataloguing upon receipt.

- Design of a print format acceptable for submission to the bibliography on archival science, which is maintained by the PAL.

- Further modification of the print format for bibliographies to more closely meet the AACR2 cataloguing standard, practised by the PAL.

- Design and implementation of a serials control module for the journals received in, and/or circulated through the Division. The next version of MINISIS is expected to have a new processor which will facilitate the introduction of such a module.

- Expansion of the classification scheme to include conservation of archival records on paper, and possibly other media, followed by inclusion of documentation on hand in the data base.

- Render the system fully operational in French and English. This would involve translation of the thesaurus terms (approximately 2,000), of the data element names, of the special message file for our application, and of the system documentation; input of all these translations into a parallel data definition and development of a sorting procedure to accommodate the alternate character set, as generated by our particular hardware configuration.

---

[9] British Standards Institution. Universal Decimal Classification. English Full Edition. UDC 77 Photography and Similar Processes. London: BSI, 1985. (BS1000(77):1985)

MINISIS allows for the segregation of the data base. The Division would use this capability to automatically extract only the records of published items for the purpose of general distribution. This would effectively result in two systems: an in-house data base of the entire collection, including confidential items acquired through the invisible college; and a separate data base, referencing published materials only. This would prevent unauthorized access to confidential information by incoming requests. As for diffusion of the public data base, PHOCUS has several possible alternatives:

- Direct dial-up to the host computer is presently possible. Infrequent requests from remote sites with a terminal could be handled this way. However, since the facility is shared with other larger applications in the PAC, this option would be better exercised through a commercial data base distributor, such as CAN/OLE or QL Systems. Other specialized networks may develop with a host computer to provide such a centralized service for all data bases in the network.

- Copies of the data base could be made on tape and sent to other sites with computers able to load them into their local DBMS. The receiving facility does not necessarily have to be using MINISIS, although such an installation would greatly simplify the procedure. This option may be more desirable at the international level, particularly where computer resources are scarce. But in North America, where online access is available to anyone with a terminal and a telephone, or a nearby university library, the dial-up method is preferable.

- Users of other online data bases will hasten to add that document delivery remains the weak link of most information networks today. A user consulting an online data base often retrieves a promising list of topical references. Unfortunately, the local source of published literature possesses very few, if any, of the titles in the list. Interlibrary loan produces slightly better returns, but the delay involved may not be worth it. However, since PHOCUS contains references only to items in the Division's possession, a copy of the source document could potentially be supplied in addition to the citation. Selection of this option would require the resolution of significant copyright, revenue and labour problems. Microfilming the entire collection seems the most promising avenue, but the demand for such a product is uncertain at this stage.

Once the data entry for the present backlog has been completed, the Division will begin exploring these various alternatives in greater detail.

APPENDIX A

(Extract from Functional Specifications and System Design Manual. Vol. 1: 2-6; 1983.)

2. FUNCTIONAL SPECIFICATIONS

2.1. Data Base Characteristics:

2.1.1. The manual system presently contains 2,500 logical records, which have already been classified, catalogued and indexed. Another 2,500 unprocessed documents are on hand. Ten different document types have been identified, each requiring a different set of fields to contain the various bibliographic elements necessary to identify the item, and to produce output that will approximate AACR2 cataloguing standards. Therefore, the test file should contain a total of 200 logical records, with at least 10 of each document type.

2.1.2. The size of the collection is growing at a rate of 1,000 documents per year. This rate is expected to continue until the total number of logical records reaches 10,000. At that point, retrospective acquisition will cease. The growth rate due to current research efforts is then expected to be 100 items per year.

2.1.3. The total size of the tuple must be variable. The average manual record presently contains 800 characters.

2.1.4. The present system has followed the ANSI Standard for Bibliographic References Z39.29-1977, as its cataloguing format. This standard was chosen due to its simplicity, and applicability to the majority of the records, which happen to be analytical citations to serial publications. In order to permit future compatibility with any automated records created by the Public Archives Library, the system must produce an output format that follows AACR2 cataloguing standards as closely as possible.

2.1.5. The use of any field in the list of data elements must be optional, thus permitting null entries and later revision of stored entries.

2.1.6. A controlled vocabulary, presently containing some 1,000 pre-coordinated terms, is the thesaurus for the descriptors, which are linked to the UDC classification numbers. Online access to this thesaurus would be desirable. Although unilingual at the moment, any online thesaurus should accommodate multiple languages.

2.1.7. An uncontrolled field must be provided for listing product names.

2.1.8. Some method must be provided for tracking the status of the document through the stages of acquisition, processing, translation and circulation.

2.1.9. An attempt must be made to provide compatibility with other MARC format based systems.

2.2. Storage Requirements:

2.2.1. Data entry must be interactive. Editing of the record must be possible at the time of the original creation of a tuple, and at later dates.

2.2.2. Some form of password protection is necessary to prevent unauthorized updating of records. Five security levels have been identified: Data Base Manager, Librarian, Operator, Staff and Other.

2.2.3. Online validation of input values is required, wherever possible, for all inverted fields. Online thesaurus control and authority files are desirable, but not mandatory.

2.2.4. A skipping function will be necessary to terminate entry for new tuples once all known data elements have been input. This would automatically enter null values in all unused data elements.

2.2.5. Multiple line scrolling is required for long data elements, so that they can be reviewed as a complete block of text.

2.2.6. Error messages must advise the operator of input errors. This should permit immediate re-entry of data.

2.2.7. Input prompts must display the previous value, if any, for entry and update of the field in question.

2.2.8. Input will be at a CRT terminal (keyboard with visual display unit).

2.2.9. Mass storage will be on hard disc of sufficient size to accommodate 10,000 logical records, plus all necessary overhead space.

2.2.10. Processing speed in the automated system must equal, if not exceed, that of the manual system, which is 20 new logical records per day.

2.2.11. Special features of the hardware, such as programmable function keys, should be used to facilitate the work of data entry.

2.3. Retrieval Requirements:

2.3.1. The search processor must have a HELP command, as most users will be infrequent operators of the system. An optional terse mode is desirable for the use of the input operator, Data Base Manager and frequent users.

2.3.2. Search options must include truncation of alphabetic data elements, combination of results from individual search keys, and arithmetic operations on numeric data elements.

2.3.3. Hard copy online output of search results must be available, at the operator's discretion.

2.3.4. The operator must be able to continue searching while an output file is being printed.

2.3.5. In order to include some subset of all data elements, the operator must be able to select from a variety of standard, pre-defined output formats.

2.3.6. The output file must be sorted according to keys specified by the operator.

2.3.7. The operator must be able to review the search results on the screen before electing to print them.

2.3.8. The command language must be English.

2.3.9. Free-text searching on non-inverted fields is desirable but not necessary.

2.3.10. The system must allow at least two operators to access the data base concurrently, for either searching or data entry functions.

2.3.11. The list of searchable fields must include:

- authors, both personal and corporate;
- titles of articles as well as journals;
- Universal Decimal Classification number;
- product names;
- language;
- country of origin;
- date of publication; and
- subject headings.

2.3.12. Some method for the collection of Performance Measurement Statistics must be provided, to include such information as:

- number of logical records created;
- number of searches made on the system; and
- time spent doing file maintenance and/or enhancement of the system.

APPENDIX B

Figures and Table

FIGURE 1: FACETED CLASSIFICATION

TITLE: "STAINS ON NEGATIVES AND PRINTS"

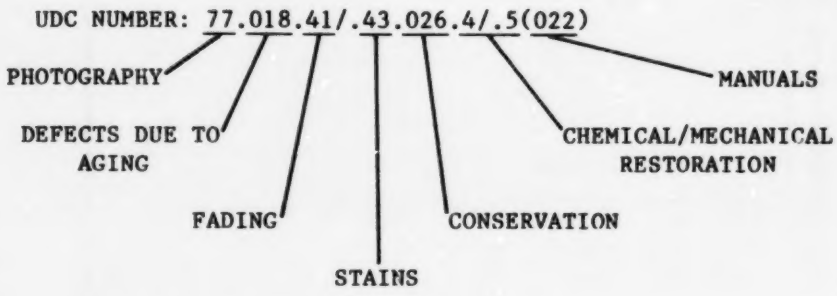


FIGURE 2: PHOCUS THESAURUS ENTRIES SHOWING RELATED TERMS

CHEMICAL RESTORATION 77.026.5

BT: CONSERVATION

NT: BLEACHING AND REDEVELOPMENT
REFIXING/REWASHING
STAIN REMOVAL
WET CLEANING

RT: MECHANICAL RESTORATION
ENHANCING PERMANENCE

ANY: RESTORATION

Stability

USE: PERMANENCE

FIGURE 3: AUTOMATIC GENERATION OF DESCRIPTORS

A) DATA ENTERED BY OPERATOR:
UDC: 77.026.1 + .4 + .5:778.1(06)

B) DATA GENERATED BY UDCEXIT:

ANALYSIS:

Main Number: 77
Auxiliary: .026.1
List: .4
List: .5
Aggregation: 778.1
Facet: (06)

POSTINGS:

Field: C350 Key: 77
Field: C350 Key: 77.026.1
Field: C350 Key: 77.026.4
Field: C350 Key: 77.026.5
Field: C350 Key: 778.1
Field: C310 Key: (06)

C) ADDITIONAL DATA STORED BY UDCEXIT:
UDCKEY(S): 77 77.026.1 77.026.4 77.026.5 778.1 (06)

D) AFTER TRANSLATION BY ONLINE THESAURUS:

77.026.1 + .4 + .5:778.1(06) = 40

AUTHOR(S): CRESPO, Carmen

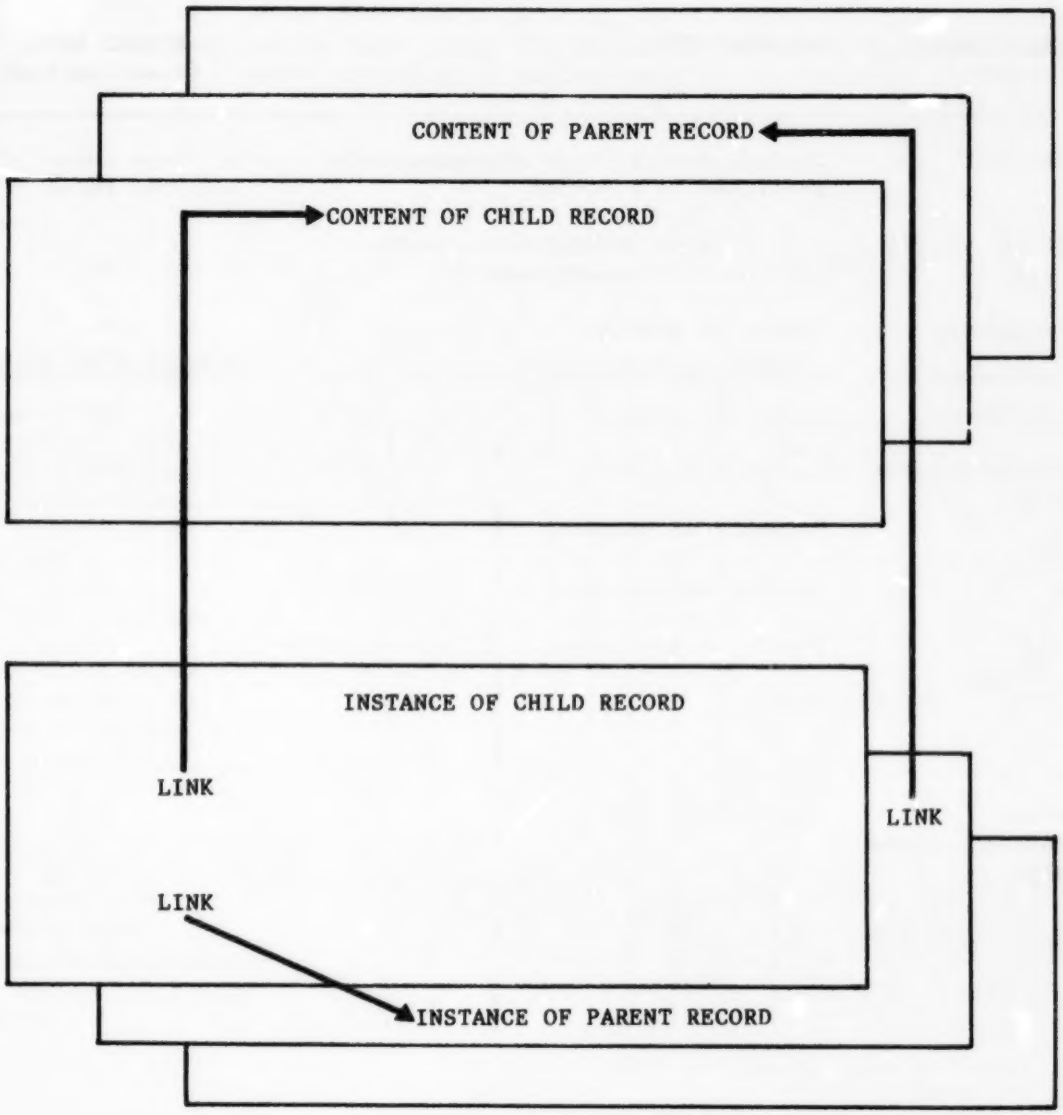
TITLE: Perfectionnement technique dans le domaine de la conservation et de la reprographie des documents d'archives. (Survey of Practice in the Field of Conservation and Reprography of Archival Documents)

SOURCE: Archives. 1978. Vol. 10 no. 1 pp. 23-37.

PRODUCT(S):

DESCRIPTOR(S): CHEMICAL RESTORATION
COPYING FOR CONSERVATION
COPYING IN GENERAL
MECHANICAL RESTORATION
PHOTOGRAPHY
ARCHIVES

FIGURE 4: RECURSIVE RECORD STRUCTURE



(CHILD CONTENT + CHILD INSTANCE) + (PARENT CONTENT + PARENT INSTANCE)

OR

PARENT CONTENT + PARENT INSTANCE

= ONE MASTER RECORD

FIGURE 5: TABLE OF RECORD TYPES

(Extract from PHOCUS Operational Documentation and User's Guide, p.31.)

BIBLIOGRAPHIC GROUP	DOCUMENT TYPE	RECORD TYPE (Function Key)
ANALYTIC	Journal Article or Conference Paper published in a serial	F1
(CHILD records only)	Book Chapter or Conference Paper published in a monograph	F2
	Patent or Standard	F3
MONOGRAPHIC	Monograph (Book)	F5
(PARENT records only)	Internal Report	F6
	Unpublished Document	F7
	Product Literature	F8
	Thesis, Student Paper, or Course Notes	F9
	Sample	F10
SERIAL (PARENT records only)	Serial Publication	F11

FIGURE 6: FIELDS PROMPTED BY FUNCTION KEYS

(Extract from PHOCUS Operational Documentation and User's Guide, pp. 28-29.)

For the next sequence in data entry, the function keys are to be used. Select the function key number according to record type.

"X" means applicable "-" means not applicable
"O" means optional

RECORD TYPES

FIELD NAME PROMPTS	CHILD F1	CHILD F2	CHILD F3	PARENT F5	PARENT F6	PARENT F7	PARENT F8	PARENT F9	PARENT F10	PARENT F11	AUTHORITY FILE
PARENT ISN	X	X	X	-	-	-	-	-	-	-	
SERIES/VOL/NO:											
SERIES NUMBER	X	-	-	-	-	-	-	-	-	-	
VOLUME NUMBER	X	O	-	-	-	-	-	-	-	-	
ISSUE NUMBER	X	-	-	-	-	-	-	-	-	-	
SOURCE DATE:											
CIRCA	X	-	X	X	X	X	X	X	X	-	
YEAR-MONTH-DAY	X	-	X	X	X	X	X	X	X	-	
SEASON	X	-	X	X	X	X	X	X	X	-	
CORPORATE SPONSOR	O	O	O	O	X	O	X	X	X	O	CORP
MEETING DETAILS	O	O	-	O	O	O	-	-	-	-	
SPECIAL NUMBER	-	-	X	-	-	-	-	-	-	-	
EDITION	-	-	-	X	O	O	O	O	-	-	
CITY OF ORIGIN	-	-	-	X	X	X	X	X	X	X	
PUBLISHER	-	-	-	X	-	-	-	-	-	O	
SERIES NOTE	O	O	-	O	O	-	X	-	-	-	
EDITOR	-	-	-	X	-	-	-	-	-	O	
DOCUMENT TYPE	-	-	-	X	X	X	X	X	X	-	
THESIS NOTE	-	-	-	-	-	-	-	X	-	-	
PUBLISHING HISTORY:											
SERIAL TITLE CODE	-	-	-	-	-	-	-	-	-	X	SERI
NOTES	O	O	O	O	O	O	O	O	O	O	

TABLE 1

(Adapted from PHOCUS Post-Implementation Report, January 9,1984, p. 1).

ITEM	PROJECTED (a)	ACTUAL
i) Development Costs:		
Design and Testing (Consultants' Fee)	$ 7,835	$16,500
Record Conversion	$ 7,206	$ 7,500
Hardware	$ 5,885	$ 5,887
Unanticipated Costs		$ 300
Total Development	$20,926	$30,187
ii) Operating Costs: (h)		
Total Operating Cost	$14,336 (f)	$10,800 (g)
iii) Disc Space and Time Spent:		
Disc Space for 1st Year	10MB (b)	8MB (e)
Development/Testing Period	4 months (c)	6 months
Conversion Period	2 months (d)	9 months

(a) Unless otherwise indicated, these figures are taken from the Revised Project Definition for an Automated System in Picture Conservation, October 8, 1982.

(b) Commercial estimate cited in Feasibility Study, January 12, 1981.

(c) As proposed by consultants who competed for the project.

(d) Progress Report, May 12, 1983.

(e) Space used as of December 1, 1983. Growth projected at 0.3 MB per month until March 31, 1985; 1 MB per year thereafter.

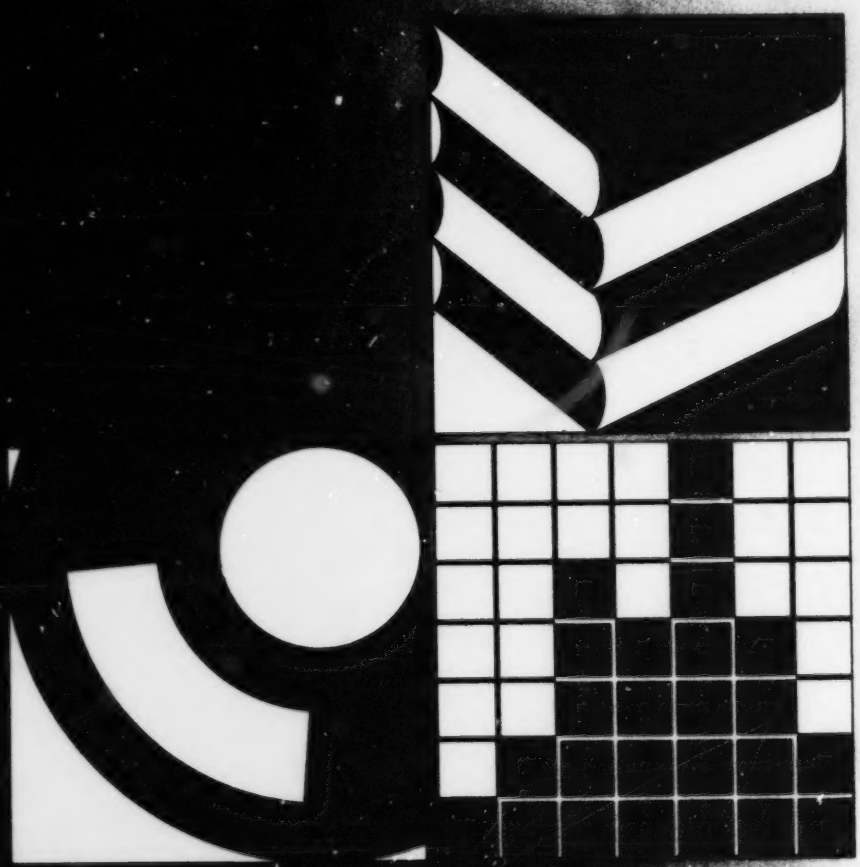
(f) Since the MINISIS user fee was not known at the time of the projection, the operating cost for the actual manual system is given instead.

(g) $10,750 budget transfer to Archives Branch as MINISIS user fee, plus $50.00 for paper for system line printer.

(h) Librarian's salary not included.

NOTES

PHOCUS



Canada

PHOCUS

BASE DE DONNÉES BIBLIOGRAPHIQUES

POUR LA

CONSERVATION DES DOCUMENTS PHOTOGRAPHIQUES

par

Klaus B. Hendriks

et

Diane Hopkins

Archives publiques Canada
Public Archives Canada

Données de catalogage avant publication (Canada)
Archives publiques Canada.
PHOCUS, base de données bibliographiques pour la conservation des documents photographiques
Texte en français et en anglais disposé tête-bêche.
Titre de la p. de t. additionnelle: PHOCUS, a bibliographic data base for the conservation of photographic materials.
Cat. MAS no SA82-3/1-1985
ISBN 0-662-53453-0
1. PHOCUS (Système d'information). 2. Systèmes d'information--Photographies--Conservation et restauration.
I. Hendriks, Klaus B. II. Hopkins, Diane. III. Titre.
IV. Titre : PHOCUS, a bibliographic data base for the conservation of photographic materials.

Z699.5.P46P8 1985 025'.0677'028 C85-097000-8F

Les Archives publiques du Canada espèrent publier dans les années à venir une série de rapports sur les techniques archivistiques, qui intéresseront sûrement un public nombreux.

© Ministre des Approvisionnements et Services Canada 1985
N° de cat. : SA82-3/1-1985
ISBN : 0-662-53453-0

TABLE DES MATIÈRES

AVANT-PROPOS v

1 INTRODUCTION 1

2 HISTORIQUE DU SYSTÈME 1

- La collection..... 1
- Objectifs du système..... 2
- Choix de MINISIS..... 3

3 ÉLABORATION DE PHOCUS 5

- Thésaurus en direct..... 6
- UDCEXIT : un programme d'analyse du numéro de classification décimale universelle..... 6
- Structure récurrente de l'enregistrement..... 6
- Touches de fonction programmables..... 7
- Le "point de vue de l'usager" pour décoder et filtrer les données..... 7
- Profils de sécurité permettant une mesure du rendement..... 8

4 CONCLUSION 8

- PHOCUS aujourd'hui..... 8
- Perfectionnements futurs..... 8

ANNEXE A : Extrait de Functional Specifications and System Design Manual..... 13

ANNEXE B : Figures et tableau..... 17

AVANT-PROPOS

La pellicule photographique traitée demeure une méthode pratiquement idéale d'entreposer des informations dans un espace réduit. Ses applications pratiques, telles les microfilms et les microfiches, les négatifs et les clichés historiques, les films de cinéma, les films aériens, les films industriels et médicaux aux rayons X, les plaques spectroscopiques, les films de duplication et de copie et autres, peuvent être en noir et blanc ou en couleurs. Les fabricants et les utilisateurs de ces types de documents s'intéressent à leur stabilité et aux méthodes de conservation.

Au cours des neuf dernières années, la Division de la restauration des peintures et estampes des Archives publiques du Canada a recueilli une quantité considérable de données publiées sur la fabrication des matériaux sensibles à la lumière, les propriétés des pellicules et supports photographiques traités, en particulier pour ce qui est de leur stabilité, leur conservation par le contrôle de l'environnement et leur restauration. Toute cette documentation a été très utile pour déterminer les facteurs modifiant la durée de vie des documents photographiques traités et pour planifier et exécuter des expériences de laboratoire visant à approfondir les connaissances de la Division dans le domaine.

Une base de données automatisée a été mise au point afin de répertorier ces riches sources d'information et d'en autoriser la consultation à partir de nombreux points d'accès. Cette base de données porte le nom de PHOCUS. L'entrée des données se poursuit, mais le système PHOCUS fonctionne déjà très bien, comme l'ont prouvé les démonstrations qui ont eu lieu aux Archives publiques du Canada, ainsi qu'à Los Angeles (Californie), grâce à une liaison téléphonique avec Ottawa, au cours de la conférence annuelle de l'American Institute for Conservation (AIC). La plus grande partie de la base de données se rapporte surtout à des documents photographiques traités, mais le système PHOCUS contient également plusieurs centaines de références à des documents connexes, en particulier des documents sur papier, ce matériau classique des bibliothèques et des archives, ainsi qu'un grand nombre de citations bibliographiques récentes se rapportant à la conservation des informations enregistrées sur support magnétique, et des documents relevant des techniques nouvelles que sont le vidéodisque et le disque optique numérique.

Le présent rapport a pour but de résumer les activités de planification et d'élaboration de la base de données PHOCUS. Comme cette base de données est unique dans le domaine de la conservation du patrimoine culturel, elle pourra éventuellement servir de modèle à d'autres bases de données portant sur des domaines connexes.

Klaus B. Hendriks
Directeur
Division de la restauration des
peintures et estampes

1 INTRODUCTION

La Division de la restauration des peintures et estampes des Archives publiques du Canada (APC) possède une collection unique d'ouvrages techniques sur la préservation des documents photographiques. Récemment, l'accès à cette collection a été automatisé grâce à la création d'un système interne d'entreposage et d'extraction des données fonctionnant à l'aide du système MINISIS. Le sigle PHOCUS désigne à la fois le système et la base de données bibliographiques qu'il soutient.

PHOCUS est l'abréviation de **PHO** tograph **C** onservation **U** niversal Decimal Classification **S** ystem (système de classification décimale universelle pour la conservation des photographies).

L'instauration de ce système présentera plusieurs avantages pour la Division, notamment :

- un contrôle de la circulation des articles et des périodiques de la collection;

- une mise à jour automatique du programme de classification et du thésaurus, accompagnée des amendements subséquents au répertoire;

- un accès supplémentaire à la documentation à partir de la date, de la source de publication, de la langue, du pays d'origine et d'autres éléments qui pourraient s'avérer nécessaires à l'avenir;

- un module de logiciel permettant de décomposer les différents éléments d'un code de classification décimale universelle afin qu'il puisse être inversé pour accélérer la recherche, avec ou sans l'aide du thésaurus.

Cette collection étant unique, les APC envisagent de mettre la base de données PHOCUS à la disposition des centres de recherche spécialisée au moyen d'une liaison directe ou en fournissant sur demande, des copies de la base de données sur bande magnétique. En ce qui concerne la distribution générale, la Division publiera des bibliographies techniques extraites de la base de données.

2 HISTORIQUE DU SYSTÈME

La collection

Au cours des années, la Division a assemblé, dans le cadre des activités de restauration des photographies, une collection de quelque 5 000 documents publiés et inédits. Il s'agit d'une vaste collection d'envergure internationale, portant sur les domaines suivants :

- l'histoire des procédés et matériaux photographiques;
- certains aspects de la science et de la technique photographiques;
- la fabrication des plaques photographiques, des pellicules et du papier d'impression;
- la photochimie et la photographie;
- la stabilité des documents photographiques traités;
- la conservation des composants.

Présentement, la plus grande partie des documents porte sur les procédés gélatino-argentiques de traitement des photographies en noir et blanc.

La collection est classée selon les principes de classification décimale universelle (CDU), et répertoriée dans un thésaurus mis au point par nos services, qui est relié au programme de classification. En utilisant le système CDU, nous avons proposé un certain nombre de mises à jour de la division analytique concernant les photographies. Ces changements ont finalement été approuvés par la Fédération Internationale de Documentation (FID), organisme responsable du système de classification, au mois de janvier 1984[1].

Le programme de classification et le thésaurus ont été mis au point afin de permettre l'accès par sujets détaillés à la collection par l'intermédiaire d'un catalogue sur fiches conventionnel.

Objectifs du système

Ce système manuel ne s'est pas avéré très efficace pour la bibliothécaire, car celle-ci ne disposait d'aucun personnel pour classer les fiches au catalogue (en moyenne 6 fiches par document répertorié). Par ailleurs, les usagers du système se sont aperçus que l'accès à la collection à partir du sujet et de l'auteur n'était pas suffisant pour retrouver certains documents dont on ne connaissait que la date ou le périodique. La production des bibliographies par matière était également lente. Lorsque l'autorisation d'automatiser le système fut obtenu, la Division put définir les objectifs suivants afin de remédier aux faiblesses du système manuel[2]:

---

[1] L'avis des modifications était publié dans Extensions and Corrections to the UDC 1983, série 12, nº 1, mai 1984, p. 19-23, (Publication FID 248/12:1).

[2] Hopkins, Diane, Is an Automated In-House Information Retrieval System a Cost-Effective Means of Providing Subject Access to a Special Collection?, London, School of Library and Information Science, University of Western Ontario, 1981, p. 10-11.

1) accepter un maximum de 10 000 enregistrements;
2) offrir un plus grand nombre de zones de recherche (15 au minimum);
3) améliorer le temps de réponse des opérations de recherche et de traitement;
4) stocker des résumés et permettre les recherches de textes libres;
5) réduire le coût de traitement des nouveaux enregistrements;
6) réduire le coût unitaire des extractions;
7) réduire le coût de production des bibliographies;
8) donner de la souplesse aux procédures de recherche et à l'organisation du système;
9) autoriser un contrôle des listes des mots-vedettes en direct;
10) rendre possible les demandes futures de données non bibliographiques;
11) offrir des produits supplémentaires sur support en papier;
12) autoriser la communication à distance.

Choix de MINISIS

La première étude de faisabilité[3] avait révélé qu'un système micro-informatique avec mémoire sur disques rigides serait le moyen le plus rentable pour réaliser les objectifs d'informatisation de la Division. Depuis, la mise sur le marché de systèmes de gestion de base de données (SGBD) pour les mini-ordinateurs rend cette option encore plus attrayante pour d'autres applications futures. À l'époque, toutefois, aucun produit particulier n'avait paru convenir. La Division avait envisagé la possibilité de participer à la mise au point conjointe d'un nouveau progiciel avec l'Institute of Machine Documentation, sous l'égide de l'UNESCO. Bien qu'elle n'ait pas poursuivi le projet, les résultats des travaux devraient être disponibles cette année[4], et méritent d'être pris en considération par tous ceux qui envisagent la création de nouvelles bases de données bibliographiques.

Finalement, la Division a eu, presque par défaut, la possibilité d'élaborer une application sur une configuration qui n'avait pas été envisagée au cours de la première étude de faisabilité. En effet, une autre division des Archives publiques du Canada venait de négocier un contrat de gérance informatique. Selon ce contrat on disposait, à un tarif fixe, de tout un mini-ordinateur dans un

---

[3] Ibid., p.35.

[4] Anon., "Information Technology Projects", General Information Programme Unisist Newsletter, vol. 11, n° 2, 1983, p. 24-25.

centre de traitement à façon, sans aucune limite applicable à la durée de liaison, au temps d'utilisation ni à l'espace disques. La capacité supplémentaire du système nous permettait de mettre au point notre application sur le même ordinateur, pour un prix bien inférieur aux tarifs normaux d'exploitation en temps partagé. Nous avions le choix entre deux SGBD déjà offerts par le système : IMAGE, le progiciel Hewlett-Packard inclut dans le système, ou MINISIS, un progiciel intégré mis au point spécialement pour les applications bibliographiques. Comme les APC étaient déjà détentrices d'une licence d'exploitation MINISIS[5], il n'était pas nécessaire de faire l'acquisition d'un exemplaire du programme. MINISIS a été choisi de préférence à IMAGE, pour les mêmes raisons qui avaient incité le groupe d'élaboration MINISIS à rejeter le système IMAGE[6]:

- possibilité de traiter des enregistrements de longueur réellement variable;
- possibilité de traiter des zones de longueur variable ou de fréquence variable;
- possibilité de traiter des sous-zones;
- possibilité d'extraire des clés de recherche intercalées dans des suites de textes;
- possibilité de traiter de longs résumés descriptifs.

Les caractéristiques mentionnées ci-dessus sont toutes essentielles pour les applications bibliographiques complexes. Voici quelques-unes des autres caractéristiques qui font que MINISIS peut généralement être utilisé avec n'importe quel système axé sur des textes écrits :

- la modularité : tous les sous-programmes sont indépendants les uns des autres, ce qui en facilite l'entretien et l'extension;

- l'indépendance des données : le logiciel reste isolé des données introduites dans le cadre de chaque application, ce qui permet d'utiliser le même logiciel pour plusieurs applications différentes sans avoir de bases de données communes;

---

5 Actuellement, MINISIS est commercialisé en Amérique du Nord exclusivement par Systemhouse Ltd.

6 Daneliuk, Faye A., Un système interactif sur mini-ordinateur pour la recherche documentaire et la gestion de bibliothèques, Ottawa, Centre de recherche pour le développement international, 1978, 16 p. (CRDI-TS14e).

- la souplesse : peut traiter des données se présentant sous pratiquement toutes les formes physiques; offre une grande variété de données de sortie; permet la redéfinition des données, même une fois que la saisie a commencé;

- la facilité d'utilisation : le langage QUERY est relativement facile à apprendre et à utiliser pour les chercheurs et usagers occasionnels;

- l'adaptation : peut accepter les données produites par d'autres systèmes informatisés; peut être utilisé par une petite organisation;

- la rentabilité : il est possible de concevoir, essayer et mettre en oeuvre de nouvelles applications en 6 à 12 mois.

La section suivante de ce rapport décrit des caractéristiques fonctionnelles plus précises de MINISIS qui sont apparues intéressantes lors de l'étape de sélection.

3 ÉLABORATION DE PHOCUS

La personne qui a assumé les fonctions du gestionnaire de la base de données a participé à part entière aux phases de conception et d'essai et sera responsable de l'entretien permanent. Ces expériences d'initiation ont peut-être prolongé la période d'élaboration, mais elles ont permis d'acquérir une solide connaissance des aspects les plus complexes du système. Il faudrait signaler que le système MINISIS n'est pas facile à utiliser, au-delà de l'entrée des données, sans une solide formation, contrairement à ce que prétendent les brochures publicitaires. En raison de sa souplesse même, les décisions à prendre sont plus nombreuses, augmentant ainsi la complexité de la conception de l'application. Cependant, toute personne possédant des compétences élémentaires en programmation et une bonne connaissance de l'application devrait être en mesure d'exercer les fonctions de gestionnaire de la base de données après avoir reçu la formation MINISIS nécessaire.

C'est à partir des objectifs précédemment mentionnés qu'ont été élaborées les caractéristiques fonctionnelles détaillées (voir l'annexe A). Ces caractéristiques ont par la suite été traduites dans une analyse organique des bases de données et des programmes d'application tenant compte des possibilités et des limites du système MINISIS. C'est à ce stade que les connaissances d'un analyste/programmeur expérimenté deviennent précieuses[7].

---

[7] M. Bill Page a fait office d'analyste/programmeur dans le cadre de ce projet qui a été confié sous contrat à Infoman Inc. (contrat nº OE283-07009).

Sans vouloir décrire ici tout le système, nous mentionnons ci-dessous certaines caractéristiques de notre application qui pourraient s'avérer utiles aux créateurs d'autres bases de données.

Thésaurus en direct

Le regroupement CDU par facettes est une sorte de code sténographique (voir la figure 1). Chaque facette correspond à un équivalent en langue anglaise[8] dans le thésaurus en direct, ce qui permet ainsi d'extraire des données en utilisant soit le numéro, soit les termes habituels. La traduction est automatique. De par sa structure, le thésaurus multilingue MINISIS autorise également les recherches relationnelles, exemple : terme générique (TG), terme spécifique (TS), terme associé (TA), groupe (GROUPE) et terme rejeté (EM). La figure 2 donne des exemples de ces relations.

UDCEXIT[9] : un programme d'analyse du numéro de classification décimale universelle

Un programme de traitement spécial appelé UDCEXIT, a été préparé pour analyser le numéro de classification décimale universelle (CDU), le décomposer en éléments et entreposer ces derniers sous la forme de descripteurs dans le fichier permanent. Cela permet de gagner du temps lors de l'introduction des données, étant donné qu'il n'est pas nécessaire de dactylographier les termes anglais équivalents. On gagne également de l'espace disques, les éléments CDU étant généralement plus courts que les descripteurs anglais. La figure 3 illustre la transformation d'un numéro CDU se présentant sous la forme d'une chaîne de caractères en une série de descripteurs anglais. Le thésaurus en direct décode automatiquement les descripteurs (clés CDU) qui sont produits et entreposés par UDCEXIT et affichent leurs équivalents anglais dans la présentation bibliographique demandée par le "point de vue de l'usager".

Structure récurrente de l'enregistrement

La figure 4 présente un graphique de la structure récurrente de l'enregistrement mis au point pour PHOCUS. Cette structure permet de réduire les risques de duplication des données au cours de l'introduction et les possibilités de redondance des données. Cette caractéristique est particulièrement utile dans le cas d'une

---

[8] Présentement, le système n'est pas opérationnel en français. Toutefois, afin d'assurer une meilleure compréhension, on a traduit la terminologie employée dans les figures.

[9] Le module de logiciel UDCEXIT sera remis à la Bibliothèque internationale des usagers, et sera distribué avec les versions futures du progiciel MINISIS.

collection regroupant de nombreuses versions d'un même document (un contenu : plusieurs versions) et lorsqu'un même document-source rassemble plusieurs éléments différents (plusieurs notices-enfants reliées à une notice-mère). Cette structure permet de décrire 10 types différents d'enregistrements (voir la figure 5). Pour les besoins de la recherche, un "point de vue de l'usager" établit une relation entre les différents composants du fichier bibliographique permanent. En conséquence, à ce niveau, les utilisateurs ne voient pas la structure sous-jacente (voir la figure 3).

Touches de fonction programmables

Les touches de fonction programmables par l'utilisateur sont une caractéristique précieuse qui permet de gagner du temps. Notre terminal Télévidéo 950 possède onze touches de fonction que l'on peut programmer pour introduire dans l'ordinateur une chaîne de caractères qui seront interprétés comme s'ils étaient dactylographiés entièrement par l'opérateur. Pour que les touches contiennent ces listes, ces dernières doivent être stockées dans un programme de l'ordinateur central, puis transférées vers les touches de fonction lorsque l'opérateur entre en communication avec MINISIS. Sinon, l'opérateur devrait programmer manuellement chacune des touches, chaque fois que le terminal est mis sous tension.

Pour que l'enregistrement bibliographique soit complet, il faut qu'un sous-ensemble différent de données vienne s'ajouter à chaque type d'enregistrement indiqué sur la figure 5. Chaque liste de zones nécessaires est stockée dans une de ces touches de fonction (voir la figure 6). L'opérateur se contente d'appuyer sur la touche qui désigne le type d'enregistrement pour faire apparaître automatiquement les zones appropriées.

Le "point de vue de l'usager" pour décoder et filtrer les données

Les systèmes de gestion de bases de données ont pour caractéristique de permettre de restructurer des données qui sont saisies une seule fois, dans des bases de données qui semblent différentes. En définissant le contenu d'un "point de vue de l'usager" pour un utilisateur, le gestionnaire de la base de données peut filtrer seulement les zones nécessaires du fichier permanent et combiner les zones codées avec les tables à consulter afin de produire des données de sortie fournissant les noms au complet plutôt que le code stocké dans le fichier permanent.

De plus, le système MINISIS offre un dispositif d'accès réservé qui peut servir à mettre de côté certains sous-ensembles d'une base de données. Par exemple, on peut diviser le fichier permanent en groupes d'enregistrements confidentiels et en groupes d'enregistrements publics.

Profils de sécurité permettant une mesure du rendement

Le système a été doté de cinq profils principaux de sécurité : gestionnaire de la base de données, bibliothécaire, enregistreur, personnel et autres. Chaque profil limite l'utilisateur à certaines bases de données ainsi qu'à certains types de fonction qu'il peut appliquer à toutes ces bases. Les noms des utilisateurs de chaque niveau permettent d'établir des statistiques de mesure du rendement, étant donné que chaque session est enregistrée par le système comptable incorporé à l'ordinateur. Par exemple, au niveau du bibliothécaire, les mentions DEVTHES et MODIFIER permettent de faire la distinction entre les activités d'élaboration du thésaurus et celles de vérification et de mise à jour des enregistrements créés par l'opérateur chargé de l'entrée des données.

4 CONCLUSION

PHOCUS aujourd'hui

La Division a achevé la conversion rétroactive d'environ 2 500 inscriptions manuelles. Le gestionnaire de la base de données s'est aperçu que le thésaurus en particulier et les types de présentation de l'impression exigeaient beaucoup d'entretien. On procède actuellement au catalogage d'un arriéré de 2 500 documents et la collection continue d'augmenter au rythme de 1 000 enregistrements par an.

Le tableau 1 résume les projections de coût et les estimations de temps ainsi que le coût et le temps réel.

Actuellement, la Division consacre peu de temps à la diffusion. Toutefois, suite à une demande de la Photo Materials Group of the American Institute for Conservation of Historic and Artistic Works, elle envisage la possibilité de publier ses plus récentes références dans le AIC Newsletter. Cette information est actuellement publiée dans les bulletins internes de la Division. Les demandes sont traitées avec la plus grande diligence possible, mais les effectifs actuels ne permettent pas, pour le moment, d'encourager les utilisateurs de l'extérieur.

Perfectionnements futurs

Notre système de classification est particulièrement utile pour ceux qui s'intéressent à la conservation photographique. La British Standards Institution prépare actuellement une édition révisée de la division analytique concernant la photographie, qui

fera état de toutes les modifications que nous avons introduites dans le programme et servira de norme internationale pour le classement de cette documentation[10].

Au sujet du système automatisé lui-même, la Division prévoit introduire les améliorations suivantes :

- Modification de la structure du module du thésaurus afin de permettre la sortie de descripteurs dans un autre langage équivalent.

- Achèvement du module afin de permettre le contrôle de la circulation de la collection. Ces travaux, entamés au cours de la phase d'élaboration, doivent faire l'objet d'une mise au point et d'essais avant la mise en service du module.

- Conception d'un format pour imprimer les informations des demandes de prêts entre bibliothèques sur des formules tel que demandé par la Bibliothèque des Archives publiques. Cela permettra d'éliminer la répétition inutile des inscriptions, entre la demande initiale d'acquisition d'un article et son catalogage sur réception.

- Conception d'un format d'impression acceptable pour les envois de données à la bibliographie en archivistique préparée par la Bibliothèque des Archives publiques.

- Modification supplémentaire du format d'impression des bibliographies afin qu'il respecte mieux les normes de catalogage AACR2 appliquées par la Bibliothèque des Archives publiques.

- Conception et mise en oeuvre d'un module de contrôle des périodiques applicable aux journaux reçus et diffusés à l'intérieur de la Division. On prévoit que la prochaine version du MINISIS sera dotée d'un nouveau processeur qui facilitera l'introduction d'un tel module.

- Étendue du système de classification à la conservation des archives sur papier et éventuellement sur d'autres supports, suivi de l'introduction dans la base de données de la documentation disponible.

- Amélioration du système afin de le rendre totalement opérationnel en français aussi bien qu'en anglais. Cela implique la traduction des termes du thésaurus (environ 2 000) et des noms des éléments de données, du fichier des messages spéciaux de notre

---

10 British Standards Institution, Universal Decimal Classification. English Full Edition. UDC 77 Photography and Similar Processes, Londres, BSI, 1985 (BSI000(77):1985).

application et de la documentation du système; l'introduction de toutes ces traductions dans une définition parallèle des données et la mise au point d'une méthode de tri pouvant accepter la série différente de caractères produite par notre configuration particulière.

Le progiciel MINISIS permet de séparer les bases de données. La Division a donc l'intention d'utiliser cette fonction pour extraire automatiquement les enregistrements se rapportant uniquement à des documents publiés et destinés à être distribués au grand public.

Par conséquent, celle-ci disposera de deux systèmes : une base de données interne de l'ensemble de la collection, y compris la documentation confidentielle acquise par l'intermédiaire des réseaux personnels d'information, ainsi qu'une autre base de données ne faisant état que des documents publiés. Cela permettrait d'éviter l'accès non autorisé à des données confidentielles. Quant à la diffusion des bases de données publiques, le système PHOCUS offre plusieurs possibilités :

- Il est présentement possible d'entrer en liaison directe avec notre ordinateur central. Les demandes peu fréquentes provenant de centres éloignés dotés d'un terminal peuvent être traitées de cette façon. Toutefois, comme la Division partage l'installation avec d'autres applications des Archives publiques, il vaudrait peut-être mieux offrir cette option par l'intermédiaire d'un distributeur commercial de base de données comme les systèmes CAN/OLE ou QL. Il se peut que d'autres réseaux spécialisés seront mis sur pied pour offrir un tel service centralisé pour toutes les bases de données du réseau.

- Il est possible d'établir des copies de la base de données sur bande magnétique et de les transmettre à des centres qui pourront les intégrer à la SGBD de leur ordinateur. Il n'est pas nécessaire que les systèmes récepteurs utilisent le MINISIS, bien que cela simplifierait grandement les opérations. Une telle option serait peut-être plus souhaitable au niveau international, en particulier dans les secteurs où les ressources informatiques sont rares. En revanche, la méthode de liaison directe est préférable en Amérique du Nord où toute personne qui possède un terminal et un téléphone ou qui peut se rendre à une bibliothèque universitaire, bénéficie d'un accès direct à la base de données de l'ordinateur central.

- Les utilisateurs d'autres bases de données en direct s'empresseront de souligner que la livraison des documents reste le point faible de la plupart des réseaux d'information actuellement en service. En effet, il arrive souvent que l'utilisateur d'une base de données en direct obtienne une liste prometteuse de références sur certains sujets, mais que le distributeur commercial d'ouvrages publiés de sa région possède très peu, voire même aucun des titres de sa liste. Les prêts entre bibliothèques

donnent des résultats un peu meilleurs, mais les délais de livraison peuvent parfois être prohibitifs. Cependant, comme PHOCUS ne contient que des références à des documents qui font partie de notre collection, nous pourrions fournir, en plus des références, un exemplaire des documents eux-mêmes. Si l'on retenait une telle option, il faudrait trouver une solution aux sérieux problèmes posés par les droits d'auteur, les recettes et la main-d'oeuvre. Le microfilmage de toute collection paraît être la solution la plus prometteuse, mais la demande d'un tel produit paraît incertaine pour le moment.

Une fois que la Division aura rattrapé le retard accumulé dans l'introduction des données, elle commencera à étudier plus en détails ces différentes possibilités.

ANNEXE A

(extrait de Functional Specifications and System Design Manual, vol. 1, p. 2-6, 1983.)

2. CARACTÉRISTIQUES FONCTIONNELLES

2.1. Caractéristiques de la base de données :

2.1.1. Le système manuel contient actuellement 2 500 enregistrements logiques qui ont déjà été classés, catalogués et répertoriés. Il reste un ensemble de 2 500 documents non traités. On a identifié dix types différents de documents, qui exigent chacun une gamme différente de zones énumérant les divers éléments bibliographiques nécessaires pour identifier l'article et produire des données de sortie qui soient assez proches des normes de catalogage AACR2. C'est pourquoi, il faudrait que le fichier d'essai contienne au total 200 enregistrements logiques, dont au moins 10 échantillons de chaque type de documents.

2.1.2. La collection augmente au rythme de 1 000 documents par an. On s'attend à ce que ce rythme se maintienne jusqu'à ce que le nombre total d'enregistrements logiques atteigne 10 000. À ce moment, l'acquisition rétroactive prendra fin. La collection continuera probablement d'augmenter, en raison des efforts de recherche en cours, à un rythme de 100 articles par an.

2.1.3. La dimension totale du tuple doit être variable. Actuellement, un enregistrement manuel moyen contient 800 caractères.

2.1.4. Dans le système actuel, le catalogage se fait selon la norme ANSI de référence bibliographique Z39.29-1977. Cette norme a été adoptée en raison de sa simplicité et de son application à la majorité des enregistrements qui sont des citations analytiques de périodiques. Afin qu'il soit compatible avec d'autres enregistrements automatisés créés par la Bibliothèque des Archives publiques, le système doit se conformer à une présentation des données de sortie se rapprochant le plus possible des normes de catalogage AACR2.

2.1.5. Il faut que l'utilisation d'une zone de la liste des éléments de données reste optionnelle afin d'autoriser les entrées vides et la révision ultérieure des entrées stockées.

2.1.6. Le thésaurus des descripteurs, qui sont reliés aux numéros de classification CDU, est un vocabulaire contrôlé, contenant actuellement environ 1 000 termes précoordonnés. L'accès en direct à ce thésaurus serait souhaitable. Bien qu'il soit unilingue en ce moment, tout thésaurus en direct devrait pouvoir accepter plusieurs langages différents.

2.1.7. La base de données doit offrir une zone non restreinte pour l'établissement d'une liste des noms de produits.

2.1.8. La base de données doit offrir une méthode quelconque de recherche du statut d'un document tout au long des diverses phases d'acquisition, de traitement, de traduction et de circulation.

2.1.9. On devra trouver un moyen de rendre le système compatible avec les autres systèmes de type MARC.

2.2. Exigences en matière de mémorisation :

2.2.1. L'entrée des données doit être interactive. La correction de l'enregistrement doit être possible au moment de la création initiale d'un tuple et ultérieurement.

2.2.2. Le système doit offrir une certaine protection à l'aide d'un mot de passe, afin d'éviter la mise à jour non autorisée de la documentation. Les cinq niveaux de sécurité suivants ont été identifiés : gestionnaire de la base de données, bibliothécaire, enregistreur, personnel et autres.

2.2.3. La validation en direct de la valeur des données d'entrée est nécessaire, chaque fois que possible, pour toutes les zones inversées. Il est souhaitable mais non pas obligatoire de disposer d'un contrôle du thésaurus et des fichiers d'autorité en direct.

2.2.4. Il sera nécessaire de disposer d'une fonction de saut pour interrompre l'introduction des données sur les nouveaux tuples une fois que tous les éléments de données connus ont été enregistrés. Cette fonction permettrait d'affecter automatiquement une valeur nulle à tous les éléments de données non utilisés.

2.2.5. On doit disposer d'une capacité de défilement multiligne pour les longs éléments de données afin de pouvoir les visualiser sous la forme de paragraphes complets.

2.2.6. Un signal d'erreur doit avertir l'opérateur des erreurs d'introduction, afin d'autoriser la réintroduction immédiate des données.

2.2.7. Des messages de guidage doivent signaler la valeur précédente, s'il en est, applicable à l'entrée et la mise à jour de la zone en question.

2.2.8. L'introduction des données se fera à un terminal à écran cathodique (clavier avec écran de visualisation).

2.2.9. La mémoire de masse aura pour support un disque rigide de capacité suffisante pour accepter 10 000 enregistrements logiques, plus l'espace nécessaire aux opérations de service.

2.2.10. La vitesse de traitement du système automatisé doit être égale, sinon supérieure à celle du système manuel qui est de 20 nouveaux enregistrements logiques par jour.

2.2.11. Il faudrait utiliser les caractéristiques spéciales de l'équipement telles que les touches de fonctions programmables pour faciliter l'introduction des données.

2.3. Exigences en matière d'extraction :

2.3.1. Le processeur de recherche doit avoir une commande SOS, étant donné que la plupart des utilisateurs ne seront pas très familiers avec le système. Il serait bon que le système possède un mode raccourci facultatif que pourrait utiliser l'opérateur chargé de l'introduction des données, le gestionnaire de la base de données et les personnes consultant fréquemment la base de données.

2.3.2. Les options de recherche doivent comprendre la troncature des éléments alphabétiques des données, la combinaison des résultats des différentes clés de recherche et les opérations arithmétiques des éléments de données numériques.

2.3.3. Le système doit pouvoir fournir, si l'opérateur le désire, une sortie en direct des résultats de recherche sur support en papier.

2.3.4. L'opérateur doit être en mesure de poursuivre la recherche pendant la production d'un imprimé.

2.3.5. L'opérateur doit pouvoir choisir parmi toute une gamme de présentations de sortie standard et prédéfinies, afin de pouvoir inclure certains sous-ensembles des éléments de données.

2.3.6. Le fichier de sortie doit être classé selon les commandes choisies par l'opérateur.

2.3.7. L'opérateur doit pouvoir visualiser les résultats de la recherche à l'écran avant de les faire imprimer.

2.3.8. La langue de commande doit être l'anglais.

2.3.9. La recherche des textes libres dans les zones non inversées est souhaitable mais pas nécessaire.

2.3.10. Le système doit autoriser l'accès de la base de données à au moins deux opérateurs en même temps, soit pour la recherche, soit pour l'introduction de données.

2.3.11. La liste des zones de recherche doit comprendre les zones suivantes :

- les auteurs, individus et sociétés;
- les titres, d'articles et de périodiques;
- le numéro de classification décimale universelle;
- les noms de produits;
- la langue;
- le pays d'origine;
- la date de publication;
- les vedettes-matière.

2.3.12. Le système doit offrir une méthode de collecte de statistiques de mesure du rendement permettant de recueillir les données suivantes :

- le nombre des enregistrements logiques créés;
- le nombre de recherches faites dans le système;
- le temps consacré à l'entretien et/ou à l'amélioration des fichiers du système.

ANNEXE B

Figures et tableau

FIGURE 1 : REGROUPEMENT PAR FACETTES*

TITRE : "TACHES SUR DES NÉGATIFS ET DES ÉPREUVES SUR PAPIER"

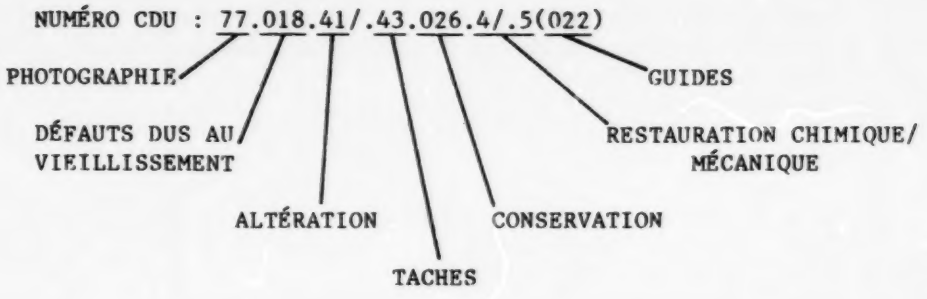


FIGURE 2 : ENTRÉES DU THÉSAURUS DE PHOCUS FAISANT ÉTAT DES TERMES ASSOCIÉS *

RESTAURATION CHIMIQUE 77.026.5

TG : CONSERVATION

TS : BLANCHIMENT ET REDÉVELOPPEMENT
REFIXAGE/RELAVAGE
ENLÈVEMENT DES TACHES
NETTOYAGE AU SOLVANT

TA : RESTAURATION MÉCANIQUE
AMÉLIORATION DE LA PERMANENCE

GROUPE : RESTAURATION

Stabilité

EM : PERMANENCE

---

* Présentement, le système n'est pas opérationnel en français. Toutefois, afin d'assurer une meilleure compréhension, on a traduit la terminologie employée dans les figures.

FIGURE 3 : PRODUCTION AUTOMATIQUE DES DESCRIPTEURS

A) DONNÉES INTRODUITES PAR L'ENREGISTREUR
CDU : 77.026.1 + .4 + .5:778.1(06)

B) DONNÉES PRODUITES PAR UDCEXIT :

ANALYSE :

Numéro principal : 77
Auxiliaire : .026.1
Liste : .4
Liste : .5
Assemblage : 778.1
Facette : (06)

INSCRIPTION :

Zone : C350 Clé : 77
Zone : C350 Clé : 77.026.1
Zone : C350 Clé : 77.026.4
Zone : C350 Clé : 77.026.5
Zone : C350 Clé : 778.1
Zone : C310 Clé : (06)

C) DONNÉES SUPPLÉMENTAIRES MISES EN MÉMOIRE PAR UDCEXIT :
CLÉ(S) CDU : 77 77.026.1 77.026.4 77.026.5 778.1 (06)

D) APRÈS TRADUCTION PAR LE THÉSAURUS EN DIRECT*:

77.026.1 + .4 + .5:778.1(06) = 40

AUTEUR(S) : CRESPO, Carmen

TITRE : Perfectionnement technique dans le domaine de la conservation et de la reprographie des documents d'archives. (Survey of Practice in the Field of Conservation and Reprography of Archival Documents)

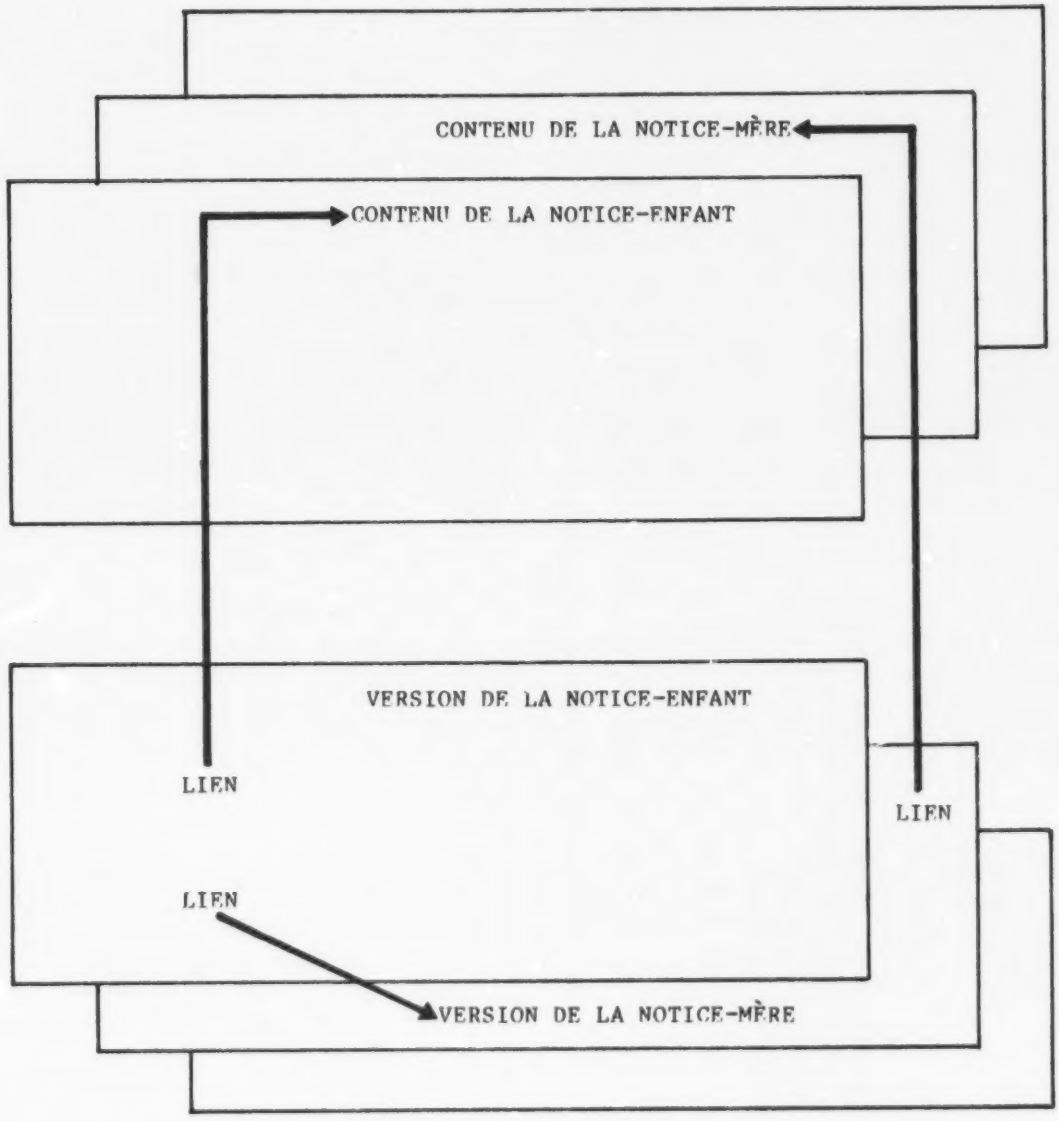
SOURCE : Archives, 1978, vol. 10 no 1, p. 23-27.

PRODUIT(S) :

DESCRIPTEUR(S) : RESTAURATION CHIMIQUE
COPIE POUR CONSERVATION
COPIE POUR USAGE GÉNÉRAL
RESTAURATION MÉCANIQUE
PHOTOGRAPHIE
ARCHIVES

* Présentement, le système n'est pas opérationnel en français. Toutefois, afin d'assurer une meilleure compréhension, on a traduit la terminologie employée dans les figures.

FIGURE 4 : STRUCTURE RÉCURRENTE DE L'ENREGISTREMENT



(CONTENU ENFANT + VERSION ENFANT) + (CONTENU MÈRE + VERSION MÈRE)

OU

CONTENU MÈRE + VERSION MÈRE

= UN ENREGISTREMENT PERMANENT

FIGURE 5 : TABLEAU DES TYPES D'ENREGISTREMENT

(extrait de Phocus Operational Documentation and User's Guide, p.31)

GROUPE BIBLIOGRAPHIQUE	TYPE DE DOCUMENT	TYPE D'ENREGISTREMENT (touche de fonction)
ANALYTIQUE	Article de journal ou compte rendu de conférences paru dans une publication en série	F1
(Notices-enfants seulement)	Chapitre d'un livre ou compte rendu d'une conférence paru dans une monographie	F2
	Brevet ou norme	F3
MONOGRAPHIQUE	Monographie (livre)	F5
(Notices-mères seulement)	Rapport interne	F6
	Document inédit	F7
	Documentation sur un produit	F8
	Thèses, essais ou notes de cours	F9
	Échantillons	F10
PUBLICATION EN SÉRIE (Notices-mères seulement)	Publication en série	F11

FIGURE 6 : ZONES INTERROGÉES PAR LES CLÉS DE FONCTION

(extrait de PHOCUS Operational Documentation and User's Guide, p. 28-29)

On utilise les touches de fonction pour la séquence suivante d'introduction des données. Il faut choisir le numéro de la touche de fonction selon le type d'enregistrement.

"X" signifie applicable "-" signifie sans objet
"O" signifie facultatif

TYPES D'ENREGISTREMENT

NOMS DES ZONES	ENFANT			MÈRE							FICHIER D'AUTORITÉ
	F1	F2	F3	F5	F6	F7	F8	F9	F10	F11	
ISN DE LA NOTICE-MÈRE	X	X	X	-	-	-	-	-	-	-	
SÉRIE/VOL/NO:											
NUMÉRO DE SÉRIE	X	-	-	-	-	-	-	-	-	-	
NUMÉRO DE VOLUME	X	O	-	-	-	-	-	-	-	-	
NUMÉRO DE PARUTION	X	-	-	-	-	-	-	-	-	-	
DATE DU DOCUMENT-SOURCE:											
CIRCA	X	-	X	X	X	X	X	X	X	-	
ANNÉE-MOIS-JOUR	X	-	X	X	X	X	X	X	X	-	
SAISON	X	-	X	X	X	X	X	X	X	-	
NOM DE COLLECTIVITÉ	O	O	O	O	X	O	X	X	X	O	CORP
DÉTAILS DE RÉUNION	O	O	-	O	O	O	-	-	-	-	
NUMÉRO SPÉCIAL	-	-	X	-	-	-	-	-	-	-	
ÉDITION	-	-	-	X	O	O	O	O	-	-	
VILLE D'ORIGINE	-	-	-	X	X	X	X	X	X	X	
ÉDITEUR	-	-	-	X	-	-	-	-	-	O	
NOTE DE SÉRIE	O	O	-	O	O	-	X	-	-	-	
DIRECTEUR DE PUBLICATION	-	-	-	X	-	-	-	-	-	O	
TYPE DE DOCUMENT	-	-	-	X	X	X	X	X	X	-	
NOTES DE THÈSE	-	-	-	-	-	-	-	X	-	-	
DÉTAILS DE PUBLICATION:											
CODE DU TITRE DE LA SÉRIE	-	-	-	-	-	-	-	-	-	X	SERI
NOTES	O	O	O	O	O	O	O	O	O	O	

TABLEAU 1

(adapté du rapport PHOCUS Post-Implementation Report, 9 janvier 1984, p. 1)

POSTE	PROJECTION (a)	CHIFFRES RÉELS
i) Coûts de mise au point :		
Conception et essai (honoraires d'experts-conseils)	7 835 $	16 500 $
Conversion des enregistrements	7 206 $	7 500 $
Matériel	5 885 $	5 887 $
Coût imprévu		300 $
Total des coûts de mise au point	20 926 $	30 187 $
ii) Coûts d'exploitation : (h)		
Total des coûts d'exploitation	14 336 $ (f)	10 800 $ (g)
iii) Espace-disques et temps utilisé :		
Espace-disques pour la première année	10 méga-octets (b)	8 méga-octets (e)
Période de mise au point/ d'essai	4 mois (c)	6 mois
Période de conversion	2 mois (d)	9 mois

(a) Sauf indication contraire, ces chiffres proviennent du document intitulé Revised Project Definition for an Automated System in Picture Conservation, 8 octobre 1982.

(b) Estimation commerciale mentionnée dans l'étude de faisabilité du 12 janvier 1981.

(c) Chiffres avancés par les experts-conseils qui ont présenté des soumissions.

(d) Progress Report, 12 mai 1983.

(e) Espace utilisé au 1er décembre 1983. On prévoit une croissance de 0,3 méga-octet par mois jusqu'au 31 mars 1985, et de 1 méga-octet par an par la suite.

(f) Étant donné que les frais d'utilisation du MINISIS n'étaient pas connus au moment de la projection, nous citons à sa place les coûts d'exploitation du système manuel actuel.

(g) Un montant de 10 750 $ à titre de transfert budgétaire à la Direction des Archives comme paiement des frais d'utilisation du MINISIS, plus 50 $ de papier pour l'imprimante en direct.

(h) Le salaire du bibliothécaire n'est pas compris.